Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT
Advances in Mathematics 188 (2004) 444–469
ADVANCES IN Mathematics
http://www.elsevier.com/locate/aim

Sandwiching random graphs: universality between random graph models

J.H. Kim[a] and V.H. Vu[b,*,1]

[a] *Microsoft Research, Microsoft Corporation, Redmond, WA 98052, USA*
[b] *Department of Mathematics, University of California, San Diego, Gilman Drive 9500, La Jolla, CA 92093, USA*

Received 9 August 2002; accepted 29 October 2003

Communicated by Laszlo Lovasz

Abstract

The goal of this paper is to establish a connection between two classical models of random graphs: the random graph $G(n,p)$ and the random regular graph $G_d(n)$. This connection appears to be very useful in deriving properties of one model from the other and explains why many graph invariants are universal. In particular, one obtains one-line proofs of several highly non-trivial and recent results on $G_d(n)$.

MSC: 05C80; 60C05; 74H50

Keywords: Random graphs; Universality

1. Introduction

The concept of random graphs is one of the central notions in modern combinatorics. Strictly speaking, the term "random graph" comprises several models of random graphs, whose treatments are usually different. Two of the most popular models, which are also the objects of this paper, are the Erdős–Rényi $G(n,p)$

*Corresponding author.
E-mail addresses: jehkim@microsoft.com (J.H. Kim), vanvu@math.ucsd.edu (V.H. Vu).
URL: http://www.math.ucsd/~vanvu.
[1] Research supported in part by an NSF Career Award, by NSF grant DMS-0200357 and by an A. Sloan fellowship.

0001-8708/$ - see front matter
doi:10.1016/j.aim.2003.10.007

model and the random regular graph model. A typical problem for these models is to show that a certain property (such as containing a complete graph of size k) holds with probability tending to 1 as n tends to infinity, given the appropriate edge density. In this case we say that the random graph has the property almost surely. Another typical problem is to determine the asymptotic behavior of a graph invariant (such as the chromatic number) at a given edge density.

The Erdős–Rényi random graph $G(n,p)$ on a set of n vertices is obtained by drawing an edge between each pair of vertices, randomly and independently, with probability p [8]. The crucial advantage of this definition is that it enables one to view samples of $G(n,p)$ as points in the probabilistic product space spanned by independent random variables. This essential fact allows us to directly apply deep and well-established results from probability theory to the study of the model. The interested readers can find many examples in the two excellent monographs by Bollobás [4] and Janson et al. [13], which are devoted to this model.

The random d-regular graph $G_d(n)$ is obtained by taking a graph uniformly at random from the set of all simple d-regular graphs on n vertices. While this definition looks simple, it, unfortunately, does not posses the powerful features of the previous one. In particular, there is no obvious relations to probability theory as in the Erdős–Rényi model. Consequently, compared to the study of Erdős–Rényi model, the study of random regular graphs relies on different techniques, usually of enumerative nature. Many of these techniques are summarized in a beautiful survey of Wormald [27].

The random regular graph model has been extensively studied in the case d is a constant and n tends to infinity and there are hundreds of papers on the subject (see, for instance, the bibliography of [27]). It turns out that arguments developed for this case frequently work for d tending slowly to infinity with n (such as $d = o(\log n)$) but collapse for larger d. For many years, results for $d \gg \log n$ ($d = n^{\Theta(1)}$, for example) were few and far between. A systematic study of this case was started only few years ago which leads to the almost complete treatments of many basic problems such as the containment of hamiltonian cycles and the asymptotic of the chromatic number (see [6,7,15,19]).

There is a remarkable fact about these new results: although the proofs are often much more complicated, all of the answers are basically the same as the answers for similar questions concerning Erdős–Rényi $G(n,p)$ model with the same edge density, which are considered rather classical results in the area. For instance, combining results from [7,19], we now know that the chromatic number of $G_d(n)$ is (a.s.) asymptotically equal to the chromatic number of $G(n,d/n)$, namely, (a.s.)

$$\chi(G_d(n) = ((1+o(1))\chi(G(n,d/n)) = (1+o(1))\frac{n\log\frac{1}{(1-d/n)}}{2\log(d)}.$$

In a way, this fact looks intuitive: For $np \gg \log n$, $G(n,p)$ is nearly np-regular by Chernoff's bound. Thus, it is sort of natural to expect that in asymptotic sense it behaves like $G_d(n)$. One would then be tempted to say that the two models are asymptotically equivalent. Obviously, it would be hugely beneficial if one could make such a statement precise. We propose the following conjecture for this purpose.

Conjecture 1 (Sandwich Conjecture). *For $d \gg \log n$, there is a joint distribution, or a coupling on (H, G_d, G) such that*

(1) *The distribution of G_d is that of the uniform random d-regular graph on n vertices. The distributions of H and G are those of the random graphs with edge probability $p_1 = \frac{d}{n}(1 - o(1))$ and $p_2 = \frac{d}{n}(1 + o(1))$, respectively. Moreover, H is a random subgraph of G with edge probability $p_H = \frac{p_1}{p_2} = 1 - o(1)$, in the sense that edges of G independently belong to H with probability p_H.*

(2) $\Pr(H \subseteq G_d) = 1 - o(1)$.

(3) $\Pr(G_d \subseteq G) = 1 - o(1)$.

Roughly speaking, our conjecture states that $G_d(n)$ can be sandwiched between two Erdős–Rényi random graphs, one with slightly smaller, the other with slightly larger, density. We strongly believe that the conjecture is true and hope that some of our arguments presented here might be sharpened to actually prove it.

The confirmation of the conjecture would reduce the study of any monotone graph property (such as being hamiltonian) or graph invariant (such as the chromatic number) of $G_d(n)$ with $d \gg \log n$ to the study of the same property or function of the more established model $G(n,p)$, where many results are already known and full machineries from probability theory can be applied. For instance, Conjecture 1 and the well-known results of Bollobás [5] and Łuczak [20] immediately yield that $\chi(G_d(n)) = (1 + o(1))\chi(G(n,p)) = (1 + o(1))\frac{n \log \frac{1}{1-p}}{2\log(np)}$, where $p = d/n$.

Throughout the paper, we say that a graph property P is (increasingly) monotone if the following holds: once $G \vdash P$, then $G' \vdash P$ for any graph G' obtained from G by adding one edge. A graph function f is (increasingly) monotone if $f(G) \leqslant f(G')$. The word "increasingly" will be omitted if there is no danger of misunderstanding.

We succeed to prove a result slightly weaker than Conjecture 1 for $\log n \ll d \ll n^{1/3}/\log^2 n$.

Theorem 2. *Let $\log n \ll d \ll n^{1/3}/\log^2 n$ and $\varphi(d)$ be any function with $(d \log n)^{1/2} \leqslant \varphi(d) \ll d$. Then there is a joint distribution, or coupling, on (H, G_d, G) such that*

(1) *The distribution of G_d is that of the uniform random d-regular graph on n vertices. The distributions of H and G are those of the Erdős–Rényi random graphs with edge probability $p_1 = \frac{d}{n}(1 - O((\frac{\log n}{d})^{1/3}))$ and $p_2 = \frac{d}{n}(1 + O(\frac{\varphi(d)}{d}))$, respectively. Moreover, H is a random subgraph of G with edge probability $p_H = \frac{p_1}{p_2} = 1 - O(\frac{\varphi(d)}{d} + (\frac{\log n}{d})^{1/3})$.*

(2) $\Pr(H \subseteq G_d) = 1 - o(1)$.

(3) $\Pr\left(\Delta(G \setminus G_d) \leqslant \dfrac{(1 + o(1))\log n}{\log(\varphi(d)/\log n)}\right) = 1 - o(1)$,

especially, for $d = n^{\nu}$ *with* $0 < \nu < 1/3$ *and* $\varphi(d) = d/\log\log n$,

$$\Pr\left(\Delta(G \backslash G_d) \leqslant \frac{1 + o(1)}{\nu}\right) = 1 - o(1),$$

where $\Delta(F)$ *is the maximum degree of* F.

The upper bound $d \leqslant n^{1/3}/\log^2 n$ is due to a technical reason (see Section 3) and we believe that this can be improved by refining certain parts of our analysis.

Part (2) of the theorem confirms part (2) of Conjecture 1 for the given range of d. Among other things, this part and the classical result of Komlós and Szemerédi [18] on the hamiltonicity of $G(n,p)$ immediately imply that $G_d(n)$ ($\log n \ll d \ll n^{1/3}/\log^2 n$) is hamiltonian. It was proved, about 10 years ago, by Frieze [9], that $G_d(n)$ is almost surely hamiltonian for $d \leqslant n^{1/5}$. But the result for all larger d was proved only very recently by Cooper et al. [6] and Krivelevich et al. [19] using highly non-trivial arguments.

Part (3) of our statement is little bit weaker than Conjecture 1 as G does not entirely contain $G_d(n)$. However, the fact that only few edges from each vertex are missing is sufficient to determined the asymptotic behavior of many graph invariants such as the chromatic number (see the last section for more details).

As shown above, Theorem 2 may yield very short proofs of known results. One can also use it to deduce new results (see Section 5.1, for examples). In particular, the theorem sheds some light on a crucial question concerning monotonicity and thresholds of the random regular graphs model (see Section 5.3).

The coupling argument developed in this paper is inspired by the coupling argument used earlier by the first author in [14] to show that the Erdős–Rényi random k-uniform hypergraph contains a perfect matching with high probability provided the expected average degree is much larger than $n^{1/(5+2/(k-1))}$. There are substantial differences between the two papers. In [14], only half of the sandwich was needed and the coupling was in the opposite direction, i.e., to derive properties of the Erdős–Rényi model from a nearly regular model.

In order to obtain the claimed coupling, we shall present a scheme which produces a triple (H, G_d, G) with the required properties. This scheme partially relies on an algorithm which generates random regular graphs with nearly uniform distribution. This algorithm is an important one and has been studied earlier by Steger and Wormald [23], but our analysis improves theirs. The core of this improvement is the new concentration results developed by the authors in [16] (see also [1], Chapter 7). The upper bound for d in Theorem 2 is required in this part of the proof. This algorithm runs in sub-quadratic time and works well in practice. (The reader is invited to check the second's author webpage for a demonstration.) A recent paper [17] contains a more delicate analysis of this algorithm.

The rest of the paper is organized as follows. In the next section, we present the algorithm generating a random regular graph and few related theorems. In Section 3, we prove these theorems by tightening Steger–Wormald's analysis using new concentration results. Next, we complete the proof of Theorem 2 in Section 4. The last section, which is much less technical, is devoted to applications and open questions. At the end of this section, we introduce a new notion of graph tolerance, which seems to be a source for a series of interesting questions.

2. The algorithm

One way to generate a random d-regular graph is to use the pairing model (see [2,3,26]: for given d and each vertex $v \in V$, let $(v,1), \dots, (v,d)$ denote the d copies of v. Let V_d denote the product set $V \times [d]$. We are going to consider graphs on V_d. An edge $\{(v,i),(w,j)\}$ of (the complete graph on) V_d with $v \neq w$ is called a copy of the edge $e = \{v,w\}$. Conversely, if an edge f in (the complete graph on) V_d is a copy of e in V, then e is called the projection of f. We call a copy of e *valid* with respect to a matching M on V_d if it shares no vertex in V_d with other edges of M and no other copy of e exists in M. When M is understood by context, we will just call it valid. An edge f in V_d is called a valid edge, or copy, if f is a valid copy of an edge e in V. The projection, denoted by $\widetilde{M}$, of a matching M on V_d is the graph on V consisting of the projections of all edges of M. A matching in V_d is called simple if its projection is a simple graph.

In [23], Steger and Wormald analyzed the following algorithm to generate a random d-regular graph.

Algorithm A. (1) *Initially $\Omega_0 = V_d$ and $M_0 = \emptyset$.*

(2) *At time t, choose two vertices i and j in Ω_t uniformly at random. If ij is not a valid edge, then repeat the random selection. If ij is an valid edge, then add it to M_t, i.e., $M_{t+1} = M_t \cup \{ij\}$, and delete i and j from Ω_t, i.e., $\Omega_{t+1} = \Omega_t \backslash \{i,j\}$. Stop if there is no valid edge.*

(3) *Set G_d be the projection of $M := M_{\text{last}}$.*

Let $\Pr_A(G)$ denote the probability that the output of Algorithm A is the graph G. In general, G_d may not be a d-regular graph, though we will show that it is with high probability if d is small. Notice that step (2) is a way to do the following task.

$(2)^*$ *Take an edge f in V_d uniformly at random among all valid edges w.r.t. M_t and set $M_{t+1} = M \cup \{f\}$. Repeat this until no valid edge can be found.*

Let $N(n,d)$ be the number of d-regular graphs on n vertices. McKay and Wormald [21] showed that for $d = o(n^{1/2})$

$$N(n,d) = \frac{(nd)!}{(\frac{nd}{2})! 2^{nd/2} (d!)^n} \exp\left(\frac{1-d^2}{4} - \frac{d^3}{12n} + O\left(\frac{d^2}{n}\right)\right). \tag{2.1}$$

If a d-regular graph is chosen uniformly at random among all $N(n,d)$ d-regular graphs, then the probability, denoted by $\Pr_u(G)$, that a d-regular graph G is chosen is $1/N(n,d)$. Steger and Wormald showed that if d is relatively small, then for most

d-regular graphs $\Pr_A$ and $\Pr_u$ are asymptotically the same, where $\Pr_A(G)$ is defined above.

Theorem 3. *If $d = o((n/\log^3 n)^{1/11})$, then for all but at most $o(N(n,d))$ d-regular graphs G*

$$\Pr_A(G) = (1 + o(1))\Pr_u(G).$$

Thus Algorithm A essentially generates a uniform random d-regular graph if $d = o((n/\log^3 n)^{1/11})$. We shall prove the following theorem, which, among other things, extends the upper bound on d in Theorem 3 to $o(n^{1/3}/\log^2 n)$.

Theorem 4. *There is a positive constant K such that if $d = o(n^{1/2})$, then for all d-regular graphs G*

$$\Pr_A(G) \geqslant (1 - o(1))\exp\left(-\frac{Kd^{3/2}\log^3 n}{n^{1/2}}\right)\Pr_u(G),$$

(*where $o(1)$ goes to* 0 *uniformly in G as n goes to* 0).

This theorem implies that for $d \ll n^{1/3}/\log^2 n$, the distribution of G generated by Algorithm A is asymptotically uniform from below, and then yields a nice coupling between $\Pr_A$ and $\Pr_u$.

Corollary 5. *If $d = o(n^{1/3}/\log^2 n)$ then for all d-regular graphs G*

$$\Pr_A(G) \geqslant (1 - o(1))\Pr_u(G).$$

Consequently, in this range of d, all but at most $o(N(n,d))$ d-regular graphs G satisfy

$$\Pr_A(G) = (1 + o(1))\Pr_u(G).$$

Corollary 6. *If $d = o(n^{1/3}/\log^2 n)$, then there is a coupling* P *on the set of pairs (H, G) of two d-regular graphs such that*

$$\Pr(H) = \Pr_A(H), \quad \mathrm{P}(G) = \Pr_u(G)$$

and

$$\Pr(H = G) = 1 - o(1).$$

We omit the rather trivial proof of Corollary 5. The proof of Corollary 6 goes as follows. For any d-regular graph F, let $p_F \coloneqq \min\{1, \Pr_u(F)/\Pr_A(F)\}$. We first generate M by Algorithm A and set $H = \widetilde{M}$. Second, let G be $\widetilde{M}$ with probability $p_{\widetilde{M}}$ and otherwise $G = F$ with probability proportional to weights

$w_F \coloneqq \Pr_A(F)(1-p_F) = \max\{\Pr_A(F) - \Pr_u(F), 0\}$. That is,

$$G = \begin{cases} \widetilde{M} & \text{with probability } p'_M \\ F & \text{with probability } (1-p_M)w_F/\sum_{F'} w_{F'} \end{cases}$$

for all d-regular graphs F. Then for any d-regular graph F

$$\begin{aligned}\Pr(G = F) &= \Pr_A(F)p_F + \sum_{F'} \Pr_A(F')(1-p_{F'})\frac{w_F}{\sum_{F''} w_{F''}} \\ &= \Pr_A(F)p_F + w_F \sum_{F'} \frac{w_{F'}}{\sum_{F''} w_{F''}} \\ &= \Pr_A(F)p_F + w_F = \Pr_u(F),\end{aligned}$$

which means that G is uniform. On the other hand, since

$$\Pr(H = F \ \& \ G = F) \geqslant \Pr_A(F)p_F = \min\{\Pr_A(F), \Pr_u(F)\}.$$

Theorem 4 gives

$$\Pr(H = G) \geqslant \sum_F \min\{\Pr_A(F), \Pr_u(F)\} \geqslant 1 - o(1),$$

provided $d = o(n^{1/3}/\log^2 n)$.

Remark. While we focus on $d = o(n^{1/3})$, Theorem 4 could be used for larger d up to $o(n^{1/2})$. An application for larger d can be found in the last section. The theorem gives only a lower bound, thus it is not at all certain that the distribution of the graph generated by the algorithm is asymptotically uniform. In Steger and Wormald [23], showed that the distribution is indeed asymptotically uniform for $d = O(n^{1/28})$. Improving the analysis developed in this paper, we are able to show that the distribution is asymptotically uniform for $d = O(n^{1/3}/\log^c n)$ for some $c > 0$ (see [17]). In other words, in Corollary 5, the inequality can be replaced by the equality if $d = O(n^{1/3}/\log^c n)$.

3. Analysis of Algorithm A—Proof of Theorem 4

It is well known and easy to check that each d-regular graph G can be projected from exactly $(d!)^n$ distinct simple perfect matchings on V_d. Let the corresponding uniform probability (over the space of all simple perfect matchings) be

$$\Pr_u(M) = \frac{1}{(d!)^n N(n,d)} = \frac{\left(\frac{nd}{2}\right)! 2^{nd/2}}{(nd)!} \exp\left(\frac{d^2-1}{4} + \frac{d^3}{12n} + O\left(\frac{d^2}{n}\right)\right),$$

and the (multiplicative standard) deviation

$$\alpha(n,d) = \exp\left(-\frac{Kd^{3/2}\log^3 n}{n^{1/2}}\right).$$

To prove Theorem 4, it suffices to show that if $d = o(n^{1/2})$, then for any simple perfect matching M

$$\begin{aligned}\Pr_A(M) &\geqslant (1-o(1))\alpha(n,d)\Pr_u(M)\\ &= \frac{(1-o(1))\alpha(n,d)\left(\frac{nd}{2}\right)!2^{nd/2}}{(nd)!}\exp\left(\frac{d^2-1}{4}+\frac{d^3}{12n}\right).\end{aligned} \tag{3.1}$$

Given a simple perfect matching M, there are $\left(\frac{nd}{2}\right)!$ ways to order the edges of M. An ordering $\mathscr{M}$ is called *good* if the probability that the edges of M appear (in the algorithm) in the same order is at least

$$(1-o(1))\frac{2^{nd/2}}{(nd)!}\exp\left(\frac{d^2-1}{4}+\frac{d^3}{12n}\right).$$

(Strictly speaking, we have to say that $\mathscr{M}$ is ϵ-good if the probability is at least

$$(1-\epsilon)\frac{2^{nd/2}}{(nd)!}\exp\left(\frac{d^2-1}{4}+\frac{d^3}{12n}\right),$$

and show that the conclusions below hold for all small enough $\epsilon > 0$. All statements below must be interpreted in this sense.) We notice that

$$\frac{2^{nd/2}}{(nd)!}\exp\left(\frac{d^2-1}{4}+\frac{d^3}{12n}\right) = \frac{(1-o(1))\exp\left(\frac{d^2-1}{4}+\frac{d^3}{12n}\right)}{\prod_{m=0}^{nd/2-1}\binom{nd-2m}{2}}. \tag{3.2}$$

It is now sufficient to show that almost every ordering $\mathscr{M}$ of the edges of M is good. For an ordering $\mathscr{M}$, we sometimes write $\mathscr{M} = (M_0, \dots, M_{dn/2})$ where M_t is the set of the first t edges of $\mathscr{M}$, especially

$$|M_t| = t, \quad M_t \subseteq M_{t+1}, \quad \text{and} \quad M_{dn/2} = M.$$

Conversely, we write $\mathscr{M} \vdash M$ if $\mathscr{M} = (M_1, \dots, M_{dn/2})$ satisfies the above conditions.

Lemma 7. *Let M be a simple perfect matching on V_d, where $d = o(n^{1/2})$. Then there are $(1-o(1))\left(\frac{dn}{2}\right)!$ distinct sequences $\mathscr{M} \vdash M$ such that*

$$\Pr_A(\mathscr{M}) \geqslant \frac{(1-o(1))\alpha(n,d)\Pr_u(M)}{\left(\frac{dn}{2}\right)!}. \tag{3.3}$$

In the remaining part of this section, we prove Lemma 7. To count the number of valid edges at time m, notice that there are $\binom{nd-2m}{2}$ ways to form an edge. However, not all edges are valid since there are edges joining two vertices from copies of the same vertex in V or being projected to the same edge as an already existing edge. The number of invalid edges of the first type is $\Delta_m^1(\mathscr{M}) = \sum_u \binom{d-d_u}{2}$ and the number of invalid edges of the second type is $\Delta_m^2(\mathcal{M}) = \sum_{u \sim_{G_m(\mathcal{M})} v} (d - d_u)(d - d_v)$, where d_u is the degree of u in $G_m(\mathcal{M})$. Set $\Delta_m(\mathcal{M}) = \Delta_m^1(\mathcal{M}) + \Delta_m^2(\mathcal{M})$, the number of valid edges is $\binom{nd-2m}{2} - \Delta_m(\mathcal{M})$. It follows that

$$\Pr_A(\mathcal{M}) = \prod_{m=0}^{nd/2-1} \frac{1}{\binom{nd-2m}{2} - \Delta_m(\mathcal{M})}. \tag{3.4}$$

Notice that $\Delta_m^1(\mathcal{M})$, $\Delta_m^2(\mathcal{M})$ and $\Delta_m(\mathcal{M})$ are actually defined with respect to the graph on V consisting of projections of the first m edges of $\mathcal{M}$. We may extend these to all graphs G on V:

$$\Delta^1(G) = \sum_u \binom{d - d_u}{2}, \quad \Delta^2(G) = \sum_{u \sim_{G(\mathcal{M})} v} (d - d_u)(d - d_v),$$

where d_u is the degree of u in G. Of course, $\Delta(G) = \Delta^1(G) + \Delta^2(G)$. In particular, $\Delta_m^1(\mathcal{M}) = \Delta^1(G_m(\mathcal{M}))$, and so on. For a graph with m edges, it is often more convenient to use $\Delta_m^1(G)$, etc. instead of just $\Delta^1(G)$ etc., since it explicitly says how many edges G has.

Set $\mu_m^1 = \frac{1}{2}(nd - 2m)^2 \frac{(d-1)}{nd}$ and $\mu_m^2 = (nd - 2m)^2 \frac{m(d-1)^2}{n^2 d^2}$. It will be useful to think of μ_m^i as a sort of expectation of Δ_m^i with respect to a random choice of $\mathcal{M}$ (this will be clear in the next section; see also Section 3.3. of [23] for a more detailed explanation). Define $\mu_m = \mu_m^1 + \mu_m^2$. A careful calculation shows that for $d = o(n^{1/2})$

$$\prod_{m=0}^{nd/2-1} \frac{\binom{nd-2m}{2}}{\binom{nd-2m}{2} - \mu_m} = \exp\left(\frac{d^2 - 1}{4} + \frac{d^3}{12n} + o(1)\right). \tag{3.5}$$

Given (3.2), (3.4) and (3.5), it is clear that to prove (3.3), we only need to show that for almost every $\mathcal{M}$

$$\prod_{m=0}^{nd/2-1} \frac{\binom{nd-2m}{2} - \mu_m}{\binom{nd-2m}{2} - \Delta_m(\mathcal{M})} = \prod_{m=0}^{nd/2-1} \left(1 + \frac{\Delta_m(\mathcal{M}) - \mu_m}{\binom{nd-2m}{2} - \Delta_m(\mathcal{M})}\right) \geqslant (1 - o(1))\alpha(n, d) \tag{3.6}$$

To prove (3.6), we show that for a typical ordering $\mathcal{M}$, the deviation $|\Delta_m - \mu_m|$ is sufficiently small, using a recent large deviation result on polynomials developed by the present authors [16]. This is the crucial difference between our analysis and the analysis in [23], which used Chernoff–Hoeffding's inequality. Our new large deviation result enables us to obtain a smaller deviation and thus improves the analysis.

3.1. Large deviation

Consider independent binary random variables $t_1, \ldots, t_n$. Let Y be a polynomial of degree k in $t_1, \ldots, t_n$ with positive coefficients. For a multi-set A of size at most k,

$\partial_A Y$ denotes the partial derivative of Y with respect to the variables t_i's *with* $i \in A$. For instance, if $Y = t_1^2 t_2^2 t_3 + t_4^5$ and $A_1 = \{1,2\}$, $A_2 = \{1,1,3\}$, then $\partial_{A_1}(F) = 4t_1 t_2 t_3$ and $\partial_{A_2}(F) = 2t_2^2$, respectively. If the set A is empty then $\partial_A Y = Y$. Finally, for all $0 \leqslant j \leqslant k$, let

$$\mathbb{E}_j(Y) = \max_{|A| \geqslant j} \mathbb{E}(\partial_A(Y))$$

The following inequality was proved by the present authors in [16] (see also [1], Chapter 7). This result was originally formulated in a different, more combinatorial, way and the version we present here is slightly stronger (see [25] for more details).

Theorem 8. *For every positive integer k there are positive constants a_k and b_k depending only on k such that the following holds. For any positive polynomial $Y = Y(t_1, \dots, t_n)$ of degree k, where t_i's are independent binary random variables*

$$\Pr\left(|Y - \mathbb{E}(Y)| \geqslant a_k \lambda^k \sqrt{\mathbb{E}_0(Y)\mathbb{E}_1(Y)}\right) \leqslant b_k e^{-\lambda/4 + (k-1)\log n}, \tag{3.7}$$

for any $\lambda > 0$.

Recall that

$$\mu_m = \frac{1}{2}(nd - 2m)^2 \left(\frac{d-1}{nd} + \frac{2(d-1)^2 m}{n^2 d^2}\right).$$

Set $p_m = 2m/nd$ and $q_m = 1 - p_m$. Let ω be a function tending to infinity arbitrarily slowly. Set $W = (nd/2)(1 - \frac{1}{\omega d^2})$ and $T_m = c \log^3 n \sqrt{nd^3 q_m^2 (d^2 q_m + 1)}$, where c is a positive constant to be determined. We say that $\mathscr{M}$ is *normal* if for every $0 \leqslant m \leqslant W$,

$$|\Delta_m - \mu_m| < T_m.$$

If $\mathscr{M}$ is not normal, then we call it *abnormal*. In this case, there is some $0 \leqslant m \leqslant W$ such that

$$|\Delta_m - \mu_m| \geqslant T_m.$$

Lemma 9. *For any fixed simple perfect matching M, $\mathscr{M}$ is almost surely normal with respect to the uniform distribution over the set of all orderings of M.*

Proof. To show that $\Pr(\mathscr{M}$ abnormal$) = o(1)$, it is more convenient to consider the following random model. Let G_{p_m} be a random subgraph of G obtained by keeping each edge of G with probability p_m, independently, where G is the graph on V consisting of projections of edges in M. It is easy to prove that with probability at least $1/nd$, G_{p_m} has exactly m edges. Therefore, to prove the lemma, it suffices to show that for each m between 0 and W

$$\Pr(|\Delta_m(G_{p_m}) - \mu_m| \geqslant T_m) \leqslant \exp(-3\log n) = o\left(\frac{1}{n^2 d^2}\right). \tag{3.8}$$

Notice that μ_m has been defined (on purpose) to be exactly the expectation of $\Delta_m(G_{p_m})$ (and μ_m^i is the expectation of $\Delta_m^i(G_{p_m})$, respectively). Consequently, (3.8) is a corollary of the following claim. □

Claim 10. *For any m between A and B*

$$\Pr(|\Delta_m^1(G_{p_m}) - \mathbb{E}(\Delta_m^1(G_{p_m}))| \geqslant T_m/2) \leqslant \frac{1}{2}\exp(-3\log n), \tag{3.9}$$

$$\Pr(|\Delta_m^2(G_{p_m}) - \mathbb{E}(\Delta_m^2(G_{p_m}))| \geqslant T_m/2) \leqslant \frac{1}{2}\exp(-3\log n). \tag{3.10}$$

Proof. Fix an m between 0 and W. In the rest of the proof, we omit the sub-index m in all relevant quantities and use a shorthand Δ^i for $\Delta^i(G_p)$. By the definition of W and q, $q \geqslant \frac{1}{\omega d^2}$. Since $d = o(n^{1/2})$, we can choose ω so that $nd^2q^2 \gg 1$. For each edge e of G, define a random variable $t_e = 1$ if e is not chosen (in G_p) and 0 otherwise. Obviously, the t_e's are i.i.d random variables with mean q. Consider

$$\Delta^1 = \sum_u \binom{(d-d_u)}{2} = \frac{1}{2}\sum_u \sum_{e \ni u, f \ni u} t_e t_f.$$

As Δ^1 is a positive polynomial in the t_e's, we are in a good position to apply Theorem 8. First, we have $\mathbb{E}(\Delta^1) \leqslant \frac{1}{2}nd^2q^2$. Moreover, for each t_e, $\partial_{t_e}\Delta^1 = \sum_{f, f\cap e \neq \emptyset} t_f$ and any partial derivative of order 2 of Δ^1 is 0 or 1. Since $nd^2q^2 \gg 1$, it follows that $\mathbb{E}_0(\Delta^1) \leqslant \frac{1}{2}nd^2q^2$ and $\mathbb{E}_1(\Delta^1) \leqslant \max_e \mathbb{E}(\sum_{f, f\cap e\neq} t_f) + 1 \leqslant 2dq + 1$. Set $\lambda = (16 + 4\epsilon)\log n$, where ϵ is a small positive constant, and apply Theorem 8. We obtain that for some constant a

$$\Pr(|\Delta^1 - \mu^1| \geqslant a\log^2 n\sqrt{(nd^2q^2)(2dq+1)}) \leqslant b_2\exp(-(3+\epsilon)\log n)) \leqslant \frac{1}{2}\exp(-3\log n).$$

Since $a\log^2 n\sqrt{(nd^2q^2)(2dq+1)} \ll T$, the first part of the claim follows.

The proof of the second part is similar. Consider the set X of all triples (e,g,f) where e,g,f are edges of G and they (in this order) form a path of length 3 in G. A short consideration shows that $\Delta^2 = \sum_{u \sim_{G_p} v}(d-d_u)(d-d_v)$ can be expressed as

$$\sum_{(e,g,f)\in X} t_e t_f(1-t_g) = Y_1 - Y_2,$$

where $Y_1 = \sum_{(e,g,f)\in X} t_e t_f$ and $Y_2 = \sum_{(e,g,f)\in X} t_e t_f t_g$. A routine calculation (similar to the one presented for Δ^1) shows that $\mathbb{E}_0(Y_1) \leqslant nd^3q^2$ and $\mathbb{E}_1(Y_1) \leqslant d^2q + 1$. Again set $\lambda = (16 + 4\epsilon)\log n$ and apply Theorem 8, we have

$$\Pr(|Y_1 - \mathbb{E}(Y_1)| \geqslant a \log^2 n \sqrt{(nd^3q^2)(d^2q+1)} \leqslant \frac{1}{4}\exp(-3\log n). \tag{3.11}$$

For Y_2, one can show that $\mathbb{E}_0(Y_2) \leqslant nd^3q^3 + 1$ and $\mathbb{E}_1(Y_2) \leqslant d^2q^2 + 1$. Notice that Y_2 has degree 3, so setting $\lambda = (20 + 4\epsilon)\log n$ gives that for some constant a

$$\Pr(|Y_2 - \mathbb{E}(Y_2)| \geqslant a \log^3 n \sqrt{(nd^3q^3+1)(d^2q^2+1)} \leqslant \frac{1}{4}\exp(-3\log n). \tag{3.12}$$

By adjusting the constant c in the definition of T, we can guarantee that for any $\frac{1}{\omega d^2} \leqslant q \leqslant 1$

$$a \log^2 n \sqrt{(nd^3q^2)(d^2q+1)} + a \log^3 n \sqrt{(nd^3q^3+1)(d^2q^2+1)} \leqslant T/2.$$

This, together with (3.11, 3.12) imply the second statement of the claim and completes our proof. $\square$

3.2. Normal is good

To complete the proof of Theorem 4, it remains to show that a normal ordering is a good ordering. This requires an elementary, but somewhat delicate computation. Consider the last product in (3.6)

$$\Pi = \prod_{m=0}^{nd/2-1} \left(1 + \frac{\Delta_m(\mathscr{M}) - \mu_m}{\binom{nd-2m}{2} - \Delta_m}\right).$$

We are going to prove that if $\mathscr{M}$ is normal, then this product is at least $(1 - o(1))\alpha(n,d)$. To this end, we assume that $\mathscr{M}$ is normal and use short hand f_m for $(1 + \frac{\Delta_m(\mathscr{M}) - \mu_m}{\binom{nd-2m}{2} - \Delta_m})$. We split Π into two parts:

$$\Pi_1 := \prod_{m=0}^{W} f_m, \quad \Pi_2 := \prod_{m=W+1}^{nd/2-1} f_m.$$

Observe that for any $m, f_m \geqslant 1 - \frac{\mu_m}{\binom{nd-2m}{2} - \mu_m}$ and $\mu_m \leqslant (nd-2m)^2(\frac{1}{n} + \frac{m}{n^2}) = o((nd - 2m)^2)$. We have

$$\Pi_2 \geqslant \prod_{m=W+1}^{nd/2-1} \left(1 - 3\left(\frac{1}{n} + \frac{m}{n^2}\right)\right) \geqslant 1 - 3 \sum_{m=W+1}^{nd/2-1} \left(\frac{1}{n} + \frac{m}{n^2}\right) = 1 - o(1),$$

due to the fact that $nd/2 - W = o(n/d)$.

Now let us consider Π_1. Since $\mathscr{M}$ is normal, $|\Delta(\mathscr{M}) - \mu_m| \leqslant T_m$. On the other, $T_m = o((nd-2m)^2)$ and $\mu_m = o((nd-2m)^2)$, so it follows that $\binom{nd-2m}{2} - \Delta_m = (1/2 - o(1))(nd-2m)^2$. Therefore,

$$\Pi_1 \geqslant \prod_{m=0}^{W} \left(1 - 3\frac{T_m}{(nd-2m)^2}\right).$$

To conclude that $\Pi_1 \geqslant (1-o(1))\alpha(n,d)$, it suffices to verify that $\sum_{m=0}^{W} \frac{T_m}{(nd-2m)^2} = O(\frac{d^{3/2}}{n^{1/2}} \log^3 n)$. Recall that $T_m = c\log^3 n\sqrt{nd^3q_m^2(d^2q_m+1)} \leqslant c\log^3 n\sqrt{nd^3q_m^2}(\sqrt{d^2q_m}+1))$, where $p_m = 2m/nd$ and $q_m = 1 - p_m$. Writing $q_m = (nd-2m)/nd$ in the upper bound of T_m and simplifying, we obtain

$$\frac{T_m}{(nd-2m)^2} \leqslant c\left(\frac{d}{n}\frac{1}{\sqrt{nd-2m}}\log^3 n + \frac{d^{1/2}}{n^{1/2}}\frac{1}{(nd-2m)}\log^3 n\right).$$

First year undergraduate calculus gives

$$\sum_{m=0}^{W} \frac{1}{\sqrt{nd-2m}} \leqslant \int_0^W \frac{dx}{\sqrt{nd-2x}} \leqslant \sqrt{nd},$$

and

$$\sum_{m=0}^{W} \frac{1}{nd-2m} \leqslant \int_{x=nd-2W}^{nd} \frac{dx}{x} \leqslant \log nd.$$

It thus follows that

$$\sum_{m=0}^{W} \frac{T_m}{(nd-2m)^2} \leqslant c\left(\frac{d^{3/2}}{n^{1/2}}\log^3 n + \frac{d^{1/2}}{n^{1/2}} 2\log^4 n\right) = c\log^3 n\frac{d^{3/2}}{n^{1/2}} + o(1), \tag{3.13}$$

completing the proof. $\square$

4. Coupling with random graphs

In this section, we first define basic random variables, and then use them to define the desired random graphs H and G as in Theorem 2 and a simple matching M on V_d constructed exactly the same way Algorithm A does. We will show that the triple $(H, \widetilde{M}, G)$ constructed this way satisfies the conclusion of Theorem 2. This is enough since

$$\Pr(\widetilde{M} = G_d) = 1 - o(1)$$

according to Corollary 6.

4.1. Basic random variables and random graphs

Let L be a Poisson random variable with mean $\lambda = dn/2 + 12\varphi(d)n$. We choose L balls and randomly color each of them B (for blue) with probability $p = (dn/2 - (300/\delta^2)n\log n)/\lambda$ and R (for red) with probability $1-p$, independently of the other balls, where $\delta = (\frac{\log n}{d})^{1/3}$. A B-ball (resp. R-ball) is a ball colored B (resp. R). For the numbers L_B, L_R of B-balls and R-balls, respectively, it is well-known that L_B and L_R are independent Poisson random variables with means $p\lambda$ and $(1-p)\lambda$, respectively.

We put each B-ball into a bin chosen uniformly at random among $\binom{n}{2}$ bins indexed by edges in the complete graph on V, and then put each R-ball the same way. A bin indexed by e is to be called the *e-bin*. Let e_t be the bin containing tth ball. Clearly, $\{e_t\}$ are i.i.d uniform random edges. The number of balls in e-bin is denoted by X_e. Similarly, $X_e^{(B)}$ and $X_e^{(R)}$ denote the numbers of B-balls and R-balls, respectively, in e-bin. It is well-known that all X_e's are mutually independent Poisson with mean $\lambda/\binom{n}{2}$, and all $X_e^{(B)}$'s and $X_e^{(R)}$'s are mutually independent Poisson with means $p\lambda/\binom{n}{2}$ and $(1-p)\lambda/\binom{n}{2}$, respectively. (One may check this by actually writing down the formula for the joint distribution for two bins.) Finally, for each ball, choose a number a from the interval $[0,1]$ uniformly at random and independently of other numbers. The number assigned to tth ball is denoted by a_t. The i.i.d. random pairs (e_t, a_t) will play a central role in the coupling.

Let G consist of all edges e_t for $t = 1, \dots, L$, or equivalently, edges e with $X_e \geqslant 1$. The graph $H = H_\delta$ consists of all edges for which the corresponding bins contain at least one B-ball with its associated number $a \geqslant \delta = (\log n/d)^{1/3}$. That is, the edge set of H is

$$\{e_t : t = 1, \dots, L_B, \quad a_t \geqslant \delta\}.$$

Clearly, $H \subseteq G$. Since the event $\{e \in G\}$ is the same as the event $\{X_e \geqslant 1\}$, and X_e's are mutually independent, we know that

$$q \coloneqq \Pr(e \in G) = \Pr(X_e \geqslant 1) = 1 - e^{-\lambda/\binom{n}{2}} = \frac{d(1 + O(\varphi(d)/d))}{n},$$

independently. Thus, G is identically distributed as the random graph $G(n, q)$. It is also clear that edges in G independently belong to H with probability $(1-\delta)p = 1 - O(\frac{\varphi(d)}{d} + (\frac{\log n}{d})^{1/3})$. So, H may be regarded as a random graph with edge probability $(1-\delta)pq = \frac{d}{n}(1 - O((\frac{\log n}{d})^{1/3}))$.

4.2. Random matching on V_d

We will generate a matching on V_d using edges e_t, $t \leqslant L$ and a few extra edges so that all edges e_t with $t \leqslant L_B$ and $a_t \geqslant \delta$ are in the matching with high probability. Recall that for $t = 1, \ldots, L$, (e_t, a_t)'s are i.i.d random pairs of a uniform random edge and a uniform random number from $[0, 1]$. We extend this sequence to an infinite sequence by taking infinite number of i.i.d. random pairs, $(e_{L+1}, a_{L+1}), (e_{L+2}, a_{L+2}), \ldots$.

A naive but tempting attempt to generate a matching using $\{(e_t, a_t)\}$ would be inductively adding a valid copy of e_{t+1} to the matching M_t uniformly at random among all valid copies of e_{t+1}. However, this construction would not generate the same matching as Algorithm A does. It turns out that a kind of admission control is required to generate a matching with the same distribution.

For a matching M of V_d, $d_M(v)$ is the number of edges in M containing a copy of v, or the degree of v in $\widetilde{M}$. The number $\bar{d}_M(v)$ of uncovered copies of v in M is $d - d_M(v)$. For each edge $e = \{v, w\}$ in V and given M_t, $M_0 = \emptyset$, define the rejection probability

$$p_t(e) = 1 - \frac{\bar{d}_{M_t}(v)\bar{d}_{M_t}(w)}{\Lambda_t^2}, \quad \text{where} \quad \Lambda_t = \max_{v \in V} \bar{d}_{M_t}(v).$$

Given M_t, we reject the edge e_{t+1} with probability $p_t(e_{t+1})$. More specifically, we reject e_{t+1} if and only if $a_{t+1} < p_t(e_{t+1})$, in which case $M_{t+1} = M_t$. If $a_{t+1} \geqslant p_t(e_{t+1})$ add a valid copy of e_{t+1} to M_t uniformly at random. If there is no valid copy of e_{t+1}, then $M_{t+1} = M_t$. For any valid edge f of M_t and its projection $\widetilde{f}$, say $\{v, w\}$,

$$\Pr(M_{t+1} = M_t \cup \{f\}) = \frac{\Pr(e_{t+1} = \widetilde{f}\,)(1 - p_t(\widetilde{f}\,))}{\bar{d}_t(v)\bar{d}_t(w)} = \binom{n}{2}^{-1} \Lambda_t^{-2}, \qquad (4.1)$$

where $\bar{d}_t$ means $\bar{d}_{M_t}$. In particular, every valid edge is equally likely to be added to M_t. Moreover, if M_t has i edges, or equivalently, there are $dn - 2i$ uncovered vertices in V_d, then since there are at most $d\Lambda_t$ forbidden vertices w for each uncovered vertex v, in the sense that the edge $\{v, w\}$ is not valid, there are at least

$$(dn - 2i)((dn - 2i) - (d+1)\Lambda_t)/2 = \left(1 - O\left(\frac{d\Lambda_t}{dn - 2i}\right)\right)\binom{dn - 2i}{2} \qquad (4.2)$$

valid edges, and an edge is added to M_t with probability

$$\left(1 - O\left(\frac{d\Lambda_t}{dn - 2i}\right)\right)\binom{dn - 2i}{2}\binom{n}{2}^{-1} \Lambda_t^{-2}. \qquad (4.3)$$

Let T_i be the first time t when $|M_t| = i$, and $M_i^* = M_{T_i}$. Then we have constructed the sequence $\mathscr{M}^* = (M_i^*)$ exactly the same way Algorithm A does (see $(2)^*$ in the definition of Algorithm A). Denote the time $S_{i+1} := T_{i+1} - T_i$ to be the time required to add an edge to M_{T_i} and $\sigma_i(v)$ to be the normalized deviation of $\bar{d}_{T_i}(v)$ from the mean, i.e., $\sigma_i(v)$ satisfies

$$\bar{d}_{T_i}(v) = (d - 2i/n)(1 + \sigma_i(v)).$$

For $\sigma_i := \max_{v \in V} \sigma_i(v)$,

$$\Lambda_{T_i} = (d - 2i/n)(1 + \sigma_i).$$

Notice that the distribution of S_{i+1} conditioned on M_{T_i} is geometrical with success probability P_i, where P_i is the probability that $|M_{T_i+1}| = |M_{T_i}| + 1$, in particular

$$P_i = \left(1 - O\left(\frac{d\Lambda_{T_i}}{dn - 2i}\right)\right)\binom{dn-2i}{2}\binom{n}{2}^{-1}\Lambda_{T_i}^{-2} = \left(1 - O\left(\frac{(1+\sigma_i)d}{n}\right)\right)\frac{1}{(1+\sigma_i)^2}.$$

That is,

$$\Pr(S_{i+1} \geqslant l) = (1 - P_i)^l.$$

Generally, if the sequence $\mathcal{M}^* = (M_i^*)$ is given, then $T_i = \sum_{j=1}^{i} S_j$ is the sum of independent geometric random variables S_j with success probability P_{j-1} depending only on M_{j-1}^*. We first prove the following lemma.

Lemma 11. *Let M be a simple perfect matching on V_d and $d = o(n^{1/3}/\log^2 n)$. Then there are $(1 - o(1))(dn/2)!$ sequences $\mathcal{M} = (M_i) \vdash M$ such that*

$$|\sigma_i(v)| \leqslant 4\left(\frac{\log n}{d - 2i/n}\right)^{1/2} \quad \forall v \in V \quad \textit{and} \quad i = 1, \ldots, dn/2 - 100n\log n. \tag{4.4}$$

And, for a given sequence $\mathcal{M}$ satisfying (4.4),

$$\Pr(T_{dn/2 - 100n\log n} > dn/2 + 10n(d\log n)^{1/2} | \mathcal{M}^* = \mathcal{M}) = o(1).$$

$$\Pr(T_{dn/2 - 10^5 n(\log n)^2/\varphi(d)} - T_{dn/2 - 100n\log n} > \varphi(d)n | \mathcal{M}^* = \mathcal{M}) = o(1).$$

Proof. Take a sequence $\mathcal{M} = (M_i)$ uniformly at random among all $(dn/2)!$ sequences. Then, for fixed i, the distribution of M_i is uniform among all collections of i edges of M and

$$\Pr(\#\text{ edges in } M_i \text{ containing } v = l) = \frac{\binom{d}{l}\binom{dn/2-d}{i-l}}{\binom{dn/2}{i}},$$

for all v. Writing

$$\frac{\binom{d}{l}\binom{dn/2-d}{i-l}}{\binom{dn/2}{i}} = \frac{\binom{d}{l}q^l(1-q)^{d-l}\binom{dn/2-d}{i-l}q^{i-l}(1-q)^{dn/2-d-i+l}}{\binom{dn/2}{i}q^i(1-q)^{dn/2-i}}$$

with $q = 2i/dn$ and taking i.i.d binomial random variables X_j, for $j = 1, \dots, dn/2$, with $\Pr(X_j = 1) = 1 - \Pr(X_j = 0) = q$, we have that

$$\frac{\binom{d}{l}\binom{dn/2-d}{i-l}}{\binom{dn/2}{i}} = \frac{\Pr(X_1 + \dots + X_d = l)\Pr(X_{d+1} + \dots + X_{dn/2} = i - l)}{\Pr(X_1 + \dots + X_{dn/2} = i)}.$$

It is easy to see

$$\Pr(X_{d+1} + \dots + X_{dn/2} = i - l) \leqslant (1 + o(1))\Pr(X_1 + \dots + X_{dn/2} = i)$$

using direct comparisons for $i = o((dn)^{1/2})$, and Stirling formula for larger i, and hence

$$\frac{\Pr(X_1 + \dots + X_d = l)\Pr(X_{d+1} + \dots + X_{dn/2} = i - l)}{\Pr(X_1 + \dots + X_{dn/2} = i)} \leqslant (1 + o(1))\Pr(X_1 + \dots + X_d = l).$$

Since the expectation of the sum $(X_1 + \dots + X_d)$ is $2i/n$, it is not very hard to check (see e.g. Appendix A of [1]) that

$$\Pr(|d - (X_1 + \dots + X_d) - (d - 2i/n)| \geqslant 4((d - 2i/n)\log n)^{1/2}) = O(n^{-4}),$$

for any $i = 1, \dots, dn/2 - 100n \log n$. (One can, for instance, set $Y - j = 1 - X_j$ and apply Chernoff's bound for $\sum_j Y_j$.) The random variable $d - (X_1 + \dots + X_d)$ is more appropriate since a bound for $\bar{d}_{T_i}(v) = (d - 2i/n)(1 + \sigma_i(v))$ is needed and $\bar{d}_{T_i}(v)$ has the same distribution as the random variable. Thus the portion of the sequences (M_i) which violate (4.4) for some $v \in V$ and i is at most $O(dn^2/n^4) = o(1)$.

For $i = dn/2 - 100n \log n$, we know that

$$T_i = \sum_{j=1}^{i} S_j$$

where S_j's are independent geometric random variables with success probability

$$P_j = \left(1 - O\left(\frac{(1 + \sigma_j)d}{n}\right)\right)\frac{1}{(1 + \sigma_j)^2} = \frac{(1 - O(d/n))}{(1 + \sigma_j)^2}$$

by (4.4). Thus

$$\mathbb{E}(T_i) = \sum_{j=1}^{i} \mathbb{E}(S_j) = \sum_{j=1}^{i} \frac{1}{P_j} = (1 + O(d/n)) \sum_{j=1}^{i} (1 + \sigma_j)^2 \leqslant i + 5(d \log n)^{1/2} n,$$

and the variance

$$V(T_i) = \sum_{j=1}^{i} V(S_j) = \sum_{j=1}^{i} \frac{1-P_j}{P_j^2} \leqslant (1 - O(d/n)) \times \sum_{j=1}^{i} (1+\sigma_j)^4 \leqslant i + 10(d \log n)^{1/2} n \leqslant dn.$$

Chebyshev's inequality gives

$$\Pr(T_i \geqslant i + 10(d\log n)^{1/2} n) \leqslant \frac{dn}{25dn^2 \log n} = o(1).$$

Let $i^* = dn/2 - 10^5 n(\log n)^2/\varphi(d)$ and $S \coloneqq T_{i^*} - T_i$. Then

$$S = \sum_{j=i+1}^{i^*} S_j$$

and, for $i \leqslant j \leqslant i^*$, (4.4) gives

$$\Lambda_{T_j} \leqslant \Lambda_{T_i} \leqslant 210 \log n,$$

and

$$P_j = \left(1 - O\left(\frac{d\Lambda_t}{dn-2j}\right)\right)\binom{dn-2i}{2}\binom{n}{2}^{-1}\Lambda_t^{-2} = (1 - O(d\varphi(d)/n))\frac{(d-2j/n)^2}{(210\log n)^2}.$$

Thus

$$E(S) = \sum_{j=i+1}^{i^*} \frac{1}{P_j} \leqslant 10^5 \log^2 n \sum_{j=i+1}^{i^*} \frac{1}{(d-2j/n)^2} \leqslant \frac{10^5 n \log^2 n}{2 \times 10^5 \log^2 n/\varphi(d)} = \varphi(d)n/2$$

and

$$V(S) = \sum_{j=i+1}^{i^*} \frac{1-P_j}{P_j^2} \leqslant 10^{10} \log^4 n \sum_{j=i+1}^{i^*} \frac{1}{(d-2j/n)^4} \leqslant \frac{10^{10} n \log^4 n}{(2\times 10^5 \log^2 n/\varphi(d))^3} \leqslant (\varphi(d))^3\, n.$$

Chebyshev's inequality yields

$$\Pr(S \geqslant \varphi(d)n) \leqslant \frac{(\varphi(d))^3 n}{(\varphi(d)n/2)^2} = o(1). \qquad \square$$

Corollary 12. *For* $d = o(n^{1/3}/\log^2 n)$,

$$\Pr\left(|\sigma_i(v)| \leqslant 4\left(\frac{\log n}{d - 2i/n}\right)^{1/2} \quad \forall v \in V, \quad i = 1, \ldots, dn/2 - 100n\log n\right) = 1 - o(1), \tag{4.5}$$

$$\Pr(T_{dn/2-100n\log n} \leqslant dn/2 + 10n(d\log n)^{1/2}) = 1 - o(1), \tag{4.6}$$

$$\Pr(T_{dn/2-10^5 n(\log n)^2/\varphi(d)} - T_{dn/2-100n\log n} \leqslant \varphi(d)n) = 1 - o(1). \tag{4.7}$$

Proof. For (4.5), Lemmas 7 and 11 implies that the probability is at least

$$\begin{aligned}
\sum_{M:SPM} \sum_{\substack{\mathcal{M} \vdash M \\ \text{with (4.4)}}} \Pr(\mathcal{M}^* = \mathcal{M}) &\geqslant \sum_{M:SPM} \sum_{\substack{\mathcal{M} \vdash M \\ \text{with (4.4),(3.3)}}} \Pr(\mathcal{M}^* = \mathcal{M}) \\
&\geqslant \sum_{M:SPM} \sum_{\substack{\mathcal{M} \vdash M \\ \text{with (4.4),(3.3)}}} \frac{(1 - o(1))\Pr_u(M)}{(dn/2)!} \\
&\geqslant \sum_{M:SPM} (1 - o(1))\Pr_u(M) = 1 - o(1),
\end{aligned}$$

where we write SPM for simple perfect matching. We prove (4.6) only. The other may be proven by the same method. For $A = dn/2 + 10n(d\log n)^{1/2}$

$$\begin{aligned}
&\Pr(T_{dn/2-100n\log n} \leqslant A) \\
&\quad \geqslant \sum_{M:SPM} \sum_{\mathcal{M} \vdash M} \Pr(\mathcal{M}^* = \mathcal{M})\Pr(T_{dn/2-200n\log n/\delta^2} \leqslant A | \mathcal{M}^* = \mathcal{M}) \\
&\quad \geqslant \sum_{M:SPM} \sum_{\substack{\mathcal{M} \vdash M \\ \text{with (4.4)}}} \Pr(\mathcal{M}^* = \mathcal{M})\Pr(T_{dn/2-200n\log n/\delta^2} \leqslant A | \mathcal{M}^* = \mathcal{M}) \\
&\quad \geqslant (1 - o(1)) \sum_{M:SPM} \sum_{\substack{\mathcal{M} \vdash M \\ \text{with (4.4)}}} \Pr(\mathcal{M}^* = \mathcal{M}) \geqslant 1 - o(1). \qquad \square
\end{aligned}$$

We are ready to prove that

$$\Pr(H \nsubseteq \widetilde{M}) = o(1) \quad \text{and} \quad \Pr\left(\Delta(\widetilde{M} \backslash G) \geqslant \frac{(1 + o(1))\log n}{\log(\varphi(d)/\log n)}\right) = o(1).$$

To show $\Pr(H \nsubseteq \widetilde{M}) = o(1)$, suppose $e_t \in H \backslash \widetilde{M}$ for some $t \leqslant L_B$. Then either $T_{dn/2-200n\log n/\delta^2} < L_B$ or there is $i \leqslant dn/2 - 200n\log n/\delta^2$ such that $p_{T_i}(e_t) > \delta$, or e_t is not valid for some $t \leqslant T_{dn/2-10n\log n/\delta^2}$. Since L_B is a Poisson random variable with mean $dn/2 - 300n\log n/\delta^2$, it is easy to check that

$$\Pr(L_B \geqslant dn/2 - 200n\log n/\delta^2) = o(1).$$

In particular, $T_{dn/2-200n\log n/\delta^2} \geqslant dn/2 - 200n\log n/\delta^2 > L_B$ with probability $1-o(1)$. Since (4.5) implies with high probability that

$$p_{T_i}(e) \leqslant 1 - \frac{(1-\min_v \sigma_i(v))^2}{(1+\max_v \sigma_i(v))^2} \leqslant 20\left(\frac{\log n}{d-2i/n}\right)^{1/2} \leqslant \delta$$

for all e and $i = 1, \ldots, dn/2 - 200n\log n/\delta^2$, the second event occurs with probability $o(1)$. For the third event, we know that e_t is not valid only if $\bar{d}_t(v) = 0$ for some $v \in e_t$ or a copy of e_t already exists. If a copy of e_t already exists, then e_t must be in $\widetilde{M}$. This is not possible since $e_t \in H \setminus \widetilde{M}$. Finally, for $i \leq dn/2 - 10n\log n/\delta^2$, (4.4) implies that $\sigma_i = o(1)$ and

$$\bar{d}_{T_i}(v) = (1+o(1))(d-2i/n) > 0.$$

Thus the third event occurs with probability $o(1)$.

Regarding $G \setminus \widetilde{M}$, recall L is a Poisson random variable with mean $dn/2 + 12\varphi(d)n$ and hence

$$\Pr(L \leqslant dn/2 + 11\varphi(d)n) = o(1).$$

On the other hand, (4.6), (4.7) and $\varphi(d) \geqslant (d\log n)^{1/2}$ yield that

$$T_{dn/2-10^5 n(\log n)^2/\varphi(d)} \leqslant dn/2 + 10(d\log n)^{1/2}n + \varphi(d)n \leqslant dn/2 + 11\varphi(d)n,$$

with probability $1-o(1)$. Thus for $i^* = dn/2 - 10^5 n(\log n)^2/\varphi(d)$

$$M_{T_{i^*}} \subseteq G.$$

with probability $1-o(1)$.

It is now enough to show that

$$\Pr\left(\bar{d}_{T_{i^*}}(v) \leqslant \frac{\left(1+\frac{2\log\log(\varphi(d)/\log n)}{\log(\varphi(d)/\log n)}\right)\log n}{\log(\varphi(d)/\log n)}, \quad \forall v\right) = 1-o(1).$$

By the same argument as in the proof of Lemma 11,

$$\Pr(\bar{d}_{T_{i^*}}(v) = l) \leqslant (1+o(1))\Pr(d - (X_1 + \cdots + X_d) = d - l),$$

where X_j's are i.i.d binomial random variable with $\Pr(X_j = 1) = 1 - \Pr(X_j = 0) = 2i^*/(dn)$. Or equivalently,

$$\Pr(\bar{d}_{T_{i^*}}(v) = l) \leqslant (1 + o(1))\Pr(Y_1 + \cdots + Y_d = l),$$

where Y_j's are i.i.d binomial random variables with $\Pr(Y_j = 1) = 1 - \Pr(Y_j = 0) = 1 - 2i^*/(dn) = 2 \times 10^5 \log^2 n/(d\varphi(d))$. For

$$l_0 = \frac{\left(1 + \frac{2\log\log(\varphi(d)/\log n)}{\log(\varphi(d)/\log n)}\right)\log n}{\log(\varphi(d)/\log n)} = \frac{(1 + o(1))\log n}{\log(\varphi(d)/\log n)},$$

we have

$$\Pr(Y_1 + \cdots + Y_d \geqslant l_0) \leqslant \binom{d}{l_0}\left(\frac{2 \times 10^5 \log^2 n}{d\varphi(d)}\right)^{l_0}$$
$$\leqslant \left(\frac{2 \times 10^5 e \log^2 n}{l_0\varphi(d)}\right)^{l_0} \leqslant \left(\frac{10^6 \log n}{\varphi(d)} \log\left(\frac{\varphi(d)}{\log n}\right)\right)^{l_0} = o(n^{-1}).$$

Therefore,

$$\Pr(\bar{d}_{T_{i^*}}(v) \geqslant l_0) = o(n^{-1}) \quad \text{and hence } \Pr(\bar{d}_{T_{i^*}}(v) \geqslant l_0 \text{ for some } v) = o(1).$$

5. Applications, remarks and questions

5.1. Few applications of the main theorem

In the introduction, we gave an example where one can use our main theorem to obtain one line proofs of known results. In this section, we present a few results which can be derived from the main theorem, but no other proofs are known.

Turán type results. Let γ be a positive real number less than 1 and H be a fixed graph. We write $G \to_\gamma H$ if G has the following Turán type property: any γ-fraction of the edges of G contains a copy of H. Let C_l denote the cycle of length l. Haxell, Kohayakawa and Łuczak [10,11] proved the following theorem.

Theorem 13. *For any $\epsilon > 0$ and fixed l, there is a constant C such that the following holds almost surely for $G = G(n, p)$ with $p \geqslant Cn^{-1+\frac{1}{l-1}}$*

(i) *$G \to_\epsilon C_l$ if l is even.*
(ii) *$G \to_{\frac{1}{2}+\epsilon} C_l$ if l is odd.*

Theorem 13 and the lower bound in the sandwich theorem imply the following result for random regular graphs.

Theorem 14. *For any $\epsilon > 0$ and fixed l, there is a constant C such that the following holds almost surely for $G = G_d(n)$ with $n^{1/3}/\log^2 n \geqslant d \geqslant Cn^{\frac{1}{l-1}}$*

(i) $G \rightarrow_{\epsilon} C_l$ *if l is even.*
(ii) $G \rightarrow_{\frac{1}{2}+\epsilon} C_l$ *if l is odd.*

The probability of not containing a fixed subgraph. Another well-known problem concerning small graphs is the problem of estimating the probability that a random graph does not contain a copy of a fixed graph H, posed by Erdős and Rényi. To start, it is known that there is a threshold $p(H)$ such that for all $p = \omega(p(H))$, $G(n,p)$ almost surely contains a copy of H [4]. Thus, for these p, the probability in question is $o(1)$. The real question is how fast $o(1)$ goes to zero with n.

Fix $p \gg p(H)$ and let $X(H)$ denote the expected number of copies of H in $G(n,p)$. Answering Erdős and Rényi's question, Janson et al. [12] showed that the probability in concern is at most $e^{-\Theta(\Phi_H)}$ where $\Phi_H = \min_{H'} X(H')$ over all subgraph H' of H with at least one edge.

The sandwich theorem implies that for those H such that

$$p(H)n \ll n^{1/3}/\log^2 n, \tag{5.1}$$

a similar bound holds for the probability that $G_d(n)$ does not contain a copy of H, for any $n^{1/3}\log^2 n \geqslant d \gg p(H)n$.

For those H that do not satisfy condition (5.1), one can have a weaker, but still exponential bound using a different argument (see [15] for details).

5.2. Beyond $n^{1/3}$

In order to prove that a property P holds (a.s.) for $G_d(n)$ with $d = o(n^{1/2})$, one could sometimes make use of Theorem 4 by showing that for such d the failure probability of P in $G(n, d/n)$ is less than $\exp(-\frac{Kd^{3/2}\log^3 n}{n^{1/2}})$.

In the introduction, we mentioned that Theorem 2 implies that $G_d(n)$ is almost surely hamiltonian for $\log n \ll d \ll n^{1/3}/\log^2 n$. By the above observation, we can extend this result to $\log n \ll d = o(n^{1/2})$. Indeed, as a consequence of Komlós–Szemerédi's result the probability that $G(n, 2\log n/n)$ fails to be hamiltonian is less than $1/2$. Thus, for any $n^{1/2} \geqslant d \gg \log n$, the probability that $G(n, d/n)$ does not contain a hamiltonian cycle is less than

$$\left(\frac{1}{2}\right)^{\frac{d}{2\log n}} \leqslant \exp\left(-\frac{Kd^{3/2}\log^3 n}{n^{1/2}}\right),$$

completing the proof.

Using the same observation, one can also extend the upper bound on d in Theorem 13. In this case, fix l and let P be the property that $G \to_{\gamma} C_l$. Let X be the minimum number of edges one needs to remove from $G(n,p)$ in order to destroy all cycles of length l. It suffices to show that the probability that $|X| \leqslant (1-\gamma)n^2p$ in $G(n,p)$ is at most $\exp\left(-\frac{Kd^{3/2}\log^3 n}{n^{1/2}}\right)$, where $d = np$. As X is a Lipschitz function, this follows easily by standard martingale arguments (see [1]).

5.3. The algorithmic model

Let $G_d^A(n)$ denote the random graph obtained by Algorithm A. The graph $G_d^A(n)$, on its own right, is an interesting model and its study has been suggested by many researchers, including Rucinsky [13] and Wormald [27]. For $d = o(n^{1/28})$, Steger and Wormald proved that the distribution of $G_d^A(n)$ is asymptotically uniform (that is $\Pr_A(G) = (1-o(1))\Pr_u(G)$ for any d-regular graph G), thus this model is basically the same as the uniform model. The proof in Section 4 actually shows that for $d \gg \log n$, $G_d^A(n)$ can be coupled with $G(n, d/n)$, providing a powerful tool for the study of $G_d^A(n)$.

Theorem 15. *Let $\log n \ll d$ and $\varphi(d)$ be any function with $(d\log n)^{1/2} \leqslant \varphi(d) \ll d$. Then there is a joint distribution, or coupling, on (H, G_d, G) such that*

(1) *The distribution of G_d is that of $G_d^A(n)$. The distributions of H and G are those of the Erdős–Rényi random graphs with edge probability $\frac{d}{n}(1 - O((\frac{\log n}{d})^{1/3}))$ and $\frac{d}{n}(1 + O(\frac{\varphi(d)}{d}))$, respectively. Moreover, H is a random subgraph of G with edge probability $p_H = 1 - O(\frac{\varphi(d)}{d} + (\frac{\log n}{d})^{1/3})$.*

(2) $\Pr(H \subseteq G_d) = 1 - o(1)$

(3) $\Pr\left(\Delta(G \backslash G_d) \leqslant \dfrac{(1+o(1))\log n}{\log(\varphi(d)/\log n)}\right) = 1 - o(1)$,

especially, for $d = n^{\nu}$ with $0 < \nu < 1/3$ and $\varphi(d) = d/\log\log n$,

$$\Pr\left(H \subseteq G_d \quad \text{and} \ \Delta(G \backslash G_d) \leqslant \frac{1+o(1)}{\nu}\right) = 1 - o(1).$$

5.4. Open questions about $\Pr_A(G)$ and $\Pr_u(G)$

It is a crucial question whether the distributions of $G_d^A(n)$ and $G_d(n)$ are asymptotically the same for all d (i.e., $\Pr_A(G) = (1+o(1))\Pr_u(G)$). This question is also important from the practical point of view as we need an efficient algorithm to simulate random regular graphs with uniform distribution. In a recent paper [17], we proved that the two distributions are asymptotically the same for $d = o(n^{1/3})$ (the proof in the current paper only shows the easier part that $\Pr_A(G) \geqslant (1-o(1))\Pr_u(G)$). The algorithm works also quite well in practice and the reader is invited to the webpage of the second author for a demonstration.

A natural question is whether the two distribution are asymptotically the same for all d. Steger and Wormald [23] conjectured that the two distributions are asymptotically the same only for $d = o(n^{1/3})$. Thus, there is a chance that one cannot extend the above mentioned result from [17].

Let us recall that in order to prove Theorem 2, it suffices to prove that for most (i.e. with the exception of a $o(1)$ fraction) d-regular graphs,

$$\Pr_A(G) \geqslant (1 - o(1))\Pr_u(G). \tag{5.2}$$

This inequality would imply that for most d-regular graphs $\Pr_A(G) \leqslant (1 + o(1))\Pr_u(G)$. We proved (5.2) for $d = o(n^{1/3}/\log^2 n)$, but may be it still holds for larger d.

For the study of several properties, it is enough to achieve a weaker result that for most d-regular graph,

$$\Pr_A(G) \geqslant c\Pr_u(G), \tag{5.3}$$

where c is a positive constant. This would imply that an event with probability $o(1)$ in $G_d^A(n)$ still has probability $o(1)$ in $G_d(n)$. It is conceivable that (5.3) holds for all $n/2 \geqslant d \gg \log n$.

5.5. The monotonicity of $G_d(n)$ and the tolerance of graphs

It is clear that the Erdős–Rényi's model is (increasingly) monotone in the following sense: if P is an (increasingly) monotone property and $G(n,p) \vdash P$ almost surely, then $G(n,p') \vdash P$ almost surely for all $p' > p$. Naturally, one may expect that the random regular model $G_d(n)$ is also monotone (i.e, if $G_d(n) \vdash P$ almost surely, then $G(n, d+1) \vdash P$ almost surely, for all d). Surprisingly, the situation is far from being clear. A principal example is again hamiltonicity. It was already proved in 1984 by Robinson and Wormald [22] that almost surely $G_{n,3}$ is hamiltonian (i.e., contains a Hamilton cycle), but the same statement for $G_{n,n/3}$ has been proved only a year ago by Krivelevich, Sudakov, Wormald and the second author [19], using a completely different argument.

One may notice that $G_d(n)$ is not monotone for $d = 1$. By definition, $G_1(n)$ is a perfect matching and thus always contains a perfect matching; on the other hand, $G_2(n)$ is a union of disjoint cycles and almost never contains a perfect matching because with probability $1 - o(1)$ at least one cycle has odd length. For constant $d \geqslant 2$, it has been shown that $G_d(n)$ is indeed monotone (see [13], Chapter 9), but the proof does not extend to larger d. For that matter, it is not even clear that if a monotone property holds (almost surely) for $G_d(n)$, then it should hold (almost surely) for $G_{d+f(d)}(n)$, for any function $f(d)$ which is bounded by a fixed degree polynomial in d.

Theorem 2 sheds some light on this critical problem. We call a property P (k,p)-tolerant if the following holds almost surely: any graph obtain from a sample of $G(n,p)$ by deleting at most k edges from each vertex has P. Furthermore, $G_d(n)$ is

T-monotone with respect to P if the following holds: $G_d(n) \vdash P$ almost surely, then $G_{d+T}(n) \vdash P$ almost surely. Theorem 2 yields the following

Corollary 16. *For all constant $0 < v < 1/3$ there is a constant $k = k(v)$ such that for any $n^v \leqslant d \leqslant n^{1/3-v}$ and $(k, d/n)$-tolerant property P, $G_d(n)$ is $kd^{2/3}\log^{1/3} n$-monotone with respect to P.*

We leave the proof to the readers as an exercise.

Given a property P and a density p, it is an interesting problem to determine the largest k so that P is (k, p)-tolerant. In a new paper [24], Sudakov and the second author call this value k the tolerance of $G(n, p)$ with respect to P and work out several estimates for k (with respect to different properties). In particular, we are able to determine the asymptotic tolerance of $G(n, p)$ with respect to hamiltonicity. One can also define and study the tolerance of deterministic graphs.

References

[1] N. Alon, J. Spencer, The Probabilistic Method, Wiley, New York, 1992.
[2] E. Bender, R. Canfield, The asymptotic number of labeled graphs with given degree sequences, J. Combin. Theory Ser. A 24 (1978) 296–307.
[3] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, European J. Combin. 1 (1980) 311–316.
[4] B. Bollobás, Random Graphs, Academic Press, New York, 1985.
[5] B. Bollobás, The chromatic number of random graphs, Combinatorica 8 (1) (1988) 49–55.
[6] C. Cooper, A. Frieze, B. Reed, Random regular graphs of non-constant degree: connectivity and Hamiltonicity, Combin. Probab. Comput. 11 (3) (2002) 249–261.
[7] C. Cooper, A. Frieze, B. Reed, O. Riordan, Random regular graphs of non-constant degree: independence and chromatic number, Combin. Probab. Comput. 11 (4) (2002) 323–341.
[8] P. Erdős, A. Rényi, On random graphs I, Publ. Math. Debrecen 6 (1959) 290–297.
[9] A. Frieze, Finding Hamiltonian cycles in spare random graphs, J. Combinatorial Theory, Ser. B 44 (1988) 230–250.
[10] P. Haxell, Y. Kohayakawa, T. Łuczak, T. Turan's extremal problem in random graphs: forbidding even cycles, J. Combin. Theory Ser. B 64 (2) (1995) 273–287.
[11] P. Haxell, Y. Kohayakawa, T. Łuczak, T. Turan's extremal problem in random graphs: forbidding odd cycles, Combinatorica 16 (1) (1996) 107–122.
[12] S. Janson, T. Łuczak, A. Ruciński, An exponential bound for the probability of nonexistence of a specified subgraph in a random graph, in: Random Graphs '87, Poznań, 1987, Wiley, Chichester, 1990, pp. 73–87.
[13] S. Janson, T. Łuczak, A. Ruciński, Random Graphs, Wiley, New York, London, 2000.
[14] J.H. Kim, Perfect Matchings in Random Uniform Hypergraphs, to appear.
[15] J.H. Kim, B. Sudakov, V.H. Vu, On the asymmetry of random graphs and random regular graphs, Random Structures Algorithms 21 (3–4) (2002) 216–224.
[16] J.H. Kim, V.H. Vu, Concentration of multi-variate polynomials and its applications, Combinatorica 20 (2000) 417–434.
[17] J.H. Kim, V.H. Vu, Generating random regular graphs, Proceeding of the Thirty-Fifth Annual ACM Symposium on Theory of Computing, (STOC 2002), pp. 213–222.
[18] J. Komlós, E. Szemerédi, Hamilton cycles in random graphs. Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdös on his 60th birthday), Vol. II, pp. 1003–1010, Colloq. Math. Soc. Janos Bolyai, Vol. 10, North-Holland, Amsterdam, 1975.

[19] M. Krivelevich, B. Sudakov, V.H. Vu, N. Wormald, Random regular graphs of high degree, Random Structures Algorithms 18 (4) (2001) 346–363.
[20] T. Łuczak, The chromatic number of random graphs, Combinatorica 11 (1) (1991) 45–54.
[21] B. McKay, N. Wormald, Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$, Combinatorica 11 (1991) 369–382.
[22] R.W. Robinson, N.C. Wormald, Almost all regular graphs are hamiltonian, Random Structures Algorithms 5 (1994) 363–374.
[23] A. Steger, N. Wormald, Generating random regular graphs quickly, Combin. Probab. Comput. 8 (4) (1999) 377–396.
[24] B. Sudakov, V.H. Vu, The tolerance of graphs, manuscript.
[25] V.H. Vu, Concentration of non-Lipschitz functions and applications, Random Structures Algorithms 20 (2002) 262–316.
[26] N.C. Wormald, The asymptotic connectivity of labelled regular graphs, J. Combin. Theory, Ser. B 31 (1881) 156–167.
[27] N.C. Wormald, Models of random regular graphs, in: J.D. Lamb, D.A. Preece (Eds.), Surveys in Combinatorics, Math. Society Lecture Notes Series, Vol. 267, Cambridge Univ. Press, Cambridge, 1999, pp. 239–298.